DOCUMENT RESUME

ED 367 639 SP 035 080

AUTHOR Bol, Linda; Strage, Amy
TITLE Relationships among Teachers' Assessment Practices and their Student Outcome and Study Skill Development Goals.
PUB DATE Nov 93
NOTE 18p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (New Orleans, LA, November 10-12, 1993).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Basic Skills; *Content Validity; Course Objectives; Educational Philosophy; Educational Practices; High Schools; Outcomes of Education; Relevance (Education); Secondary School Teachers; Skill Development; Student Educational Objectives; Student Evaluation; *Study Skills; Teacher Attitudes; *Test Items; *Thinking Skills

ABSTRACT

This study was conducted to explore the relationship between teachers' stated instructional objectives, students' study skill deficiencies, and the kinds of assessment items found on tests and other course documents. High school biology teachers (N=10) were interviewed individually about their teaching philosophies and practices. Responses were aligned with the types of items that appear on their tests and on other course materials. Teachers' student learning goals and the specific type of study skills they reported students to be deficient in were categorized. Test and practice items were rated on level of processing (whether the item required basic knowledge, integration, or application) and item format (recognition or recall). Overall, teachers wanted their students to develop a general interest in and understanding of the subject area and its real world applications. They also wanted their students to develop higher-order study skills by interpreting information, managing their time and effort, and thinking critically. But their assessment practices did not support these goals. On average, over half of both test and practice items required only basic knowledge, while almost none required application. Nearly two-thirds of test questions were recognition items. (Contains 24 references.) (LL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Relationships among Teachers' Assessment Practices and their Student Outcome and Study Skill Development Goals

Linda Bol, Memphis State University,
and
Amy Strage, San Jose State University

ED 367 639

This study examined the relationships between (1) teachers' student learning outcome goals and their students' study skill deficiencies and (2) their assessment practices. Ten high school biology teachers were interviewed individually about their teaching philosophies and practices. Additionally, all course documents were analyzed. Teachers' student learning goals and the specific types of study skills they reported students to be deficient in were categorized. Test and practice items were rated on level of processing (whether the item required basic knowledge, integration, or application) and item format (recognition or recall). Overall, teachers wanted their students to develop a general interest, understanding of the subject area and its real-world applications. They also wanted their students to develop higher-order study skills by interpreting information, managing their time and effort, and thinking critically. But their assessment practices did not support these goals. On average, over half of the items (52% of test items, 53% of practice items) required only basic knowledge, while almost none required application (5% of test items, 4% of practice items). Nearly two-thirds (65%) of test items were recognition items.

There has long been widespread consensus among educational researchers, practitioners and policy makers that ideally, curriculum and instruction incorporate critical thinking and problem-solving skills (e.g., Bereiter & Scardamalia, 1987; Berliner, 1992; Biggs & Collis, 1982; Bloom, Englehart, Frust, Hill, & Krathwhohl, 1956; Newmann, 1990; Peterson, 1988; Prawat, 1993). The instructional goal to progress beyond teaching basic content as a series of disjointed facts that students are required to memorize in rote fashion is hardly a matter a contention. Though educators advocate higher-order sorts of cognitive skills and strategies for instruction, there may be little correspondence between these stated goals and actual teaching practices (McCaslin & Good, 1992; Semb & Spencer, 1976). Educators may verbalize their allegiance to promoting critical thinking, problem-solving and other kinds of higher-order thinking strategies in the classroom, but there may not be a close alignment between stated instructional goals and the course characteristics that demand and support student engagement in these more sophisticated thinking skills.

SP035080

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. Bol

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Perhaps the most salient example of the misalignment between higher-order instructional goals and classroom practices is the types of test items that students encounter on classroom or standardized achievement tests. Lofty instructional goals that encourage understanding and critical thinking are undermined by test items emphasizing recognition of factual details that are inconsistent with these goals (Linn, 1990). There is ample evidence to suggest that teacher-developed classroom tests contain an overabundance of items characterized as low-level in terms of the knowledge products or thinking skills required (Crooks, 1988; Fleming & Chambers, 1983; Haertel, 1986). In a more recent analysis of standardized tests and textbook tests used nationwide in the assessment of math and science for grades 4 through 12, Madaus and his colleagues reported that "state and district tests, and textbook tests, mutually reinforce an emphasis on low-level thinking and knowledge" (Madaus, Maxwell West, Harmon, Lomax, & Viator, 1992, p.18). The findings that tests predominately include items that reflect low-level skills and thinking is not unique to elementary grades but may also be characteristic of classroom tests even at the college level. In a study at one midwestern university, Semb and Spencer (1976) first asked college instructors to estimate the number of complex tasks (e.g. problem solving) and recall tasks (e.g. definitions) on their exams, and then compared their estimates to the actual number of complex versus recall tasks that actually appeared on their exams. Although the instructors estimated that approximately 32 percent of their exam items were complex, only about 9 percent of their items actually met the definition of a complex task. "The results of the study indicate that even though most instructors do have complex educational tasks as course goals, they do not, in fact, have a high percentage of such tasks programmed into their tests" (p.121). Overall, research suggests that classroom assessment is characterized by low-level types of items, and that instructors' achievement goals may be not reflected in the kinds of tests that they administer to their students.

The importance of tests to learning and studying cannot be overstated. A substantial percentage of students' time, both inside and outside of the classroom, is devoted to preparing for and taking tests, and there is empirical evidence to suggest that the type of test item that students anticipate is related to the kinds of study activities students will engage in as well as student

achievement (Crooks, 1988; Doyle, 1983). For example, student expectations about test format (e.g., recognition versus recall) has been shown to influence student study activities and achievement (Crooks, 1988; Meyer, 1936; Thomas, Bol, Warkentin, Wilson, Strage, & Rohwer, in press). In addition to test item format, student expectations about the cognitive complexity or demand of test items may be communicated by the presence of adjunct questions (Hamaker, 1986; Martin & Saljo, 1976), specific learning objectives (Rothkopf & Billington, 1975), advance organizers (Mayer, 1979), and the kinds of practice items or supports that students encounter on other course materials in preparation for classroom tests (Bol & Thomas, 1991; Thomas, et al, in press; Thomas, Bol and Warkentin, 1991). Each of these characteristics of courses or instructional materials has been shown to influence student study strategies and achievement.

Even when test items appear to be more cognitively challenging, instructors may engage in teaching and assessment practices that effectively reduce the demand associated with the test item. These types of practices have been termed "compensations" (Strage, Tyler, Thomas, & Rohwer, 1987), and one example of a compensation is when students have seen the actual item in advance of the test. Exposure to items in practice that were identical to items that subsequently appeared on the test (identity items) has been linked to student engagement in less productive types of study activities (Thomas et al., in press). Therefore, the goals of instruction that call for critical thinking and more sophisticated study strategies may not be realized if instructors include a high percentage of identity items on their tests.

The purpose of the present study is to explore the relationships between teachers stated instructional objectives, the student study skills they identified as in need of improvement, and the kinds of assessment items found on their tests and other course documents. The results of detailed interviews with 10 high school biology teachers were aligned with the types of items that appear on their test and on other course materials. We adopted a descriptive approach in which we collected a great deal of information from a small sample of teachers. The intent was not to generalize to all high school biology courses but instead to provide an in-depth look at teachers' perceptions about what they wanted their students to achieve in their courses, the kinds of study skills students need

to acquire and their classroom assessment practices. More specifically, our study addresses the following five research questions: (1) What were the teachers' general achievement goals for their students? (2) What kinds of study skills did the teachers identify that students needed to acquire? (3) What kinds of items appeared on teachers' classroom tests and practice materials? (4) Did teachers accurately judge whether students have seen the actual test item prior to the test? (5) Are the teachers' test and practice items consistent with their goals and the types of study skills they want their students to acquire?

METHOD

Participants

Participants were 10 high school biology teachers from the greater San Francisco Bay Area. The teachers taught one or more sections of a general biology course in one of seven different senior high schools. The grade level of students enrolled in these courses ranged from Grade 9 to Grade 12, but the large majority of students were in the 10th grade. Teachers were selected from both urban and suburban sites in order to reflect the variation in the socio-economic and ethnographic make-up of Northern California.

Procedure

The first phase of the study consisted of extensive, structured interviews with the teachers on their instructional goals and practices, the content of their courses, and characteristics of their students and schools. For the present study, the focus was on teachers' responses to three interview questions. The first item assessed over-riding achievement goals by asking teachers what they wanted students to achieve from their courses. The second asked teachers about the kinds of study skills that their students needed to work on. The third item was designed to measure teachers perceptions about the practice of exposing students to test items prior to the administration of the test. Teachers were asked to estimate the percentage of test items that students had seen or practiced in advance of the test (termed identity items).

The open-ended responses to the first two interview questions were categorized on the basis of similarity of responses or consistent themes that emerged across teachers. Both authors independently categorized the original responses with 92 percent agreement on the first item and 90 percent agreement on the second item.

In the second phase of the study, all course documents (e.g., tests, quizzes, texts, homeworks, worksheets, study guides) distributed to students were collected and analyzed for a single instructional unit, genetics. The genetics unit was selected because it was covered by all teachers and because the subject area was one that required knowledge of facts and basic principles and could be readily applied to real-world examples.

Each test and practice item on all course documents was rated for level of processing (whether the item required basic knowledge, integration, or extension), item format (recognition or recall), and identity (whether the test item appeared in identical form on any of the course materials). The course materials for each teacher were rated separately by two project staff members, and the inter-rater reliabilities ranged from .89 to .92.

RESULTS

Because the study consisted of a detailed examination of the perceptions of teachers and the types of items that appeared on their course materials, simple descriptive statistics were used to analyze the results. The first section presents the achievement goals of teachers and the kinds of study skills they think that students need to develop based on their responses to the interview. This section is followed by an analysis of all items that appeared on the teachers' course documents. Finally, the teachers perceptions of their practice of including identity items were compared with the actual number of identity items that were found on their unit tests.

Achievement Goals and Study Skills

The five categories of achievement goals described by teachers are provided in Table 1. Because teachers' open-ended responses often fell into more than one category, the numbers and percentages are based on the number of different goals and not the number of teachers. The two most frequently cited achievement goals were that students gain an appreciation and understanding of living things (25 percent of responses) and that students develop learning and study skills that provided them with the tools to answer questions (25 percent). An interest and enthusiasm for biology and science and an awareness of how science is applied to real-world issues were also identified as important achievement goals (20 percent of responses, respectively). Knowledge of specific content in biology was mentioned by only two teachers and comprised 10 percent of the responses.

Table 1. Teachers' achievement goals for students in their Biology courses.

Goals	N*	%
1. An appreciation and understanding of living things.	5	25
2. Learning and study skills (tools to answer questions).	5	25
3. An interest and enthusiasm for biology and the sciences.	4	20
4. An awareness of real-word issues (applications of science).	4	20
5. Knowledge of specific content in biology.	2	10

*Teachers gave more than one response.

Overall, the teachers described similar kinds of achievement goals that primarily represented global, higher order objectives that emphasized learning skills, motivation, understanding, and the application of knowledge. The only exception to this pattern of responses was the knowledge of specific content that was endorsed by a smaller percentage of teachers.

The teachers perceptions of the kinds of study skills that students needed to develop were more varied, resulting in seven different categories of responses. The number and percentage of responses by type of study skill is provided in Table 2. The two study skills most frequently identified were the ability to distinguish between important and unimportant information and the ability to integrate information (19 percent of responses, respectively). Teachers also commonly cited time and effort management as well as the ability to summarize and interpret information as areas that students needed to improve (14 percent, respectively). Critical thinking and problem-solving was described by two teachers, and another two teachers stated that students needed to move beyond simply rote memorization. Test-taking, note-taking, and listening skills represented three single responses that were categorized as "Other" skills.

Table 2. Teachers' perception of the kinds of study skills that students need to develop.

Study Skills	N*	%
1. Identification of important versus unimportant information	4	19
2. Integration of information	4	19
3. Time and effort management	3	14
4. Summarization and interpretation	3	14
5. Critical thinking and problem-solving	2	10
6. Engaging in processes other than rote memorization	2	10
7. Other (test-taking, note-taking, and listening skills)	3	14

*Teachers gave more than one response.

With a few exceptions, teachers identified more sophisticated study skills that called for discriminating between important and less important content; summarizing, interpreting, and integrating information; and critical thinking. Self-management skills, time and effort, management might were also areas that students needed to improve. It is noteworthy that two teachers mentioned that students need to engage in processes that require more than simple rote memorization.

Types of Items on Teachers' Tests and other Course Materials

All items on the teachers' tests and other course documents were rated on the level of processing required and on item format. The classification of these items was done separately for the teachers' test items and for the items that appeared on all other course documents.

Table 3 presents the percentage of teachers' test items classified as encoding, integration, or extension for each of the ten teachers as well as the average percentage of items by type across teachers. The overall averages indicate that over half of the test items (52 percent) were ones that required encoding or basic knowledge of the information. However, there was a good deal of variation in the percentage of encoding items by individual teacher. For example, one teachers' test had 91 percent of her items classified encoding, while another teacher had only 7 percent of her items classified as encoding. The average percentage of test items that required integration was 44 percent, but again variation between teachers was observed. An examination of the percentages of test items that demanded extension or application shows that very few of teachers included these types of items on their unit tests. Four of the teachers required no extension items, and the highest percentage of extension items that appeared on one teachers' test was only 14 percent. The average percentage of extension items across teachers was 5 percent.

Table 3. Classification of the demand of teachers' unit test items.

Teacher	Percent by Item Type		
	Encoding	Integration	Extension
1	66	23	11
2	56	39	5
3	56	39	5
4	57	36	7
5	53	47	0
6	26	71	3
7	66	34	0
8	38	62	0
9	91	9	0
10	7	79	14
Average	52	44	5

Average percentages by level of processing for the items that appeared on all other course materials were remarkably similar to those found on the teachers' tests (See Table 4). The averages were 53 percent for encoding, 43 percent for integration, and 4 percent for extension which were almost identical to the mean values obtained for the test items (52 percent, 44 percent, and 5 percent, respectively). Although there was variation between teachers, there appears to be a close correspondence between the level the processing required by the test items and those required on other course materials. Students seem to be exposed to the kinds of items in practice, with respect to level of processing, that subsequently appeared on their tests. Unfortunately, items that demanded basic knowledge of the information (encoding) were much more common than items that required application (extension) across tests and other course assignments.

Table 4. Classification of the items appearing on all other course documents.

Teacher	Percent by Item Type		
	Encoding	Integration	Extension
1	71	25	4
2	67	33	1
3	65	34	1
4	47	45	8
5	69	28	3
6	70	28	2
7	9	79	12
8	48	51	1
9	36	60	4
10	46	49	4
Average	53	43	4

Turning to item format or the production requirement of test items, it was found that most of the teachers' test items required recognition (true/false, or multiple choice) rather than recall (fill-ins, short answer, or essay), but this overall pattern was reversed when items on all other course materials was examined (See Table 5). The average percentage of recognition items on the test was 65 percent, while the average percentage of recognition items in practice was 31 percent, indicating that the production demand was much lower on tests compared to the production demand found on

all other items. Three of the teachers' tests contained all recognition items, and all but three teachers had more than half of their test items classified as recognition. In contrast, all but two of teachers had more than half of their practice items classified as demanding recall.

Table 5. The percentage of recognition versus recall items on test and practice items.

Teacher	Test Items		Practice Items	
	Recognition	Recall	Recognition	Recall
1	71	29	0	100
2	100	0	69	31
3	100	0	66	34
4	62	38	1	99
5	24	76	16	84
6	44	56	39	61
7	69	31	0	100
8	84	16	25	75
9	100	0	52	48
10	0	100	41	59
Average	65	35	31	69

In summary, the teachers' test and practice items were not very challenging in terms of the level of processing demanded, particularly when the percentages of items that required extension were considered. The teachers either had no extension items on their tests and on other practice documents or had very small percentages of items that required application of the material. The production demand of test items was also not very challenging across teachers, with most teachers relying on recognition items to evaluate students.

Perceptions about Identity Items and their Actual Use

One interview item asked teachers to estimate the number of items that students had seen in identical form prior to the test, and these estimates were compared to the actual percentage of identity items that appeared on their unit tests (See Table 6). Though the average percentages indicate that teachers tended to underestimate their practice of exposing students to items that they had seen previously (16 versus 23 percent), the comparison of perceptions and testing practices

was much more informative when the results for individual teachers were examined. For instance, two teachers responded that there were no identity items on their tests, when in fact 49 percent of their test items were ones that students had previously encountered on other course materials. Another teacher estimated that only 2 percent of her test items would be considered as identity items, while an analysis of her test revealed that 64 percent of the items were identity items. An opposite pattern was observed for Teacher 10, who said that students had seen 95 percent of the items prior to the test, but who really did not include any identity items on his exam.

Table 6. The comparison teachers' perceptions of whether students had been exposed to identical items and the actual percentage of identical items on teachers' tests.

Teacher	Teachers' Perception	Percentage on Tests
1	20	0
2	0	49
3	0	49
4	10	10
5	2	64
6	20	0
7	15	6
8	0	8
9	0	40
10	95	0
Average	16	23

This comparison between perceptions about the kind of items that appear on teachers tests and the teachers' actual testing practices suggests that teachers do not accurately judge whether students have seen the items in advance of the test. The findings for the percentage of identity items on teachers tests is perhaps even more disconcerting. On four of the teachers' exams, 40 percent of more of the items were ones that students had previously seen in identical form. The practice of providing students with identity items is assumed to reduce the challenge of the test because it effectively provides students an answer that they simply memorize in preparation for the exam and thereby reduces the cognitive demands associated with the test. Not only is little

challenge associated with the teachers' test in their original form but the challenge is even further diminished because the students are exposed to the test items in advance of the test.

DISCUSSION

Though the teachers' instructional goals were exemplary and would seem to promote higher-order thinking skills, the types of items found on their tests and on other course documents do not reinforce these more global achievement goals. In other words, there does not appear to be a close correspondence between what teachers' want students to achieve in their courses and the kinds of practice and test items students are exposed to in their courses. On the average, only five percent of teachers' test items and only four percent of teachers' practice items demanded extension or application of the material, and the majority of test items were recognition versus recall. The findings that teachers' tests contained predominately low-level items in terms of the knowledge or thinking skills demanded is consistent with other research studies that analyzed the types of items appearing on classroom, standardized, and textbook tests (Fleming & Chambers, 1983, Haertel, 1986, Madaus, 1992). In fact, the percentage of extension test items (5 percent) observed in this study was remarkably similar to the percentage of problem-solving items (9 percent) reported by Semb and Spencer (1976) in their analysis of test items appearing on college instructors' exams.

The misalignment between achievement goals and assessment practices may also help explain why students may not develop the study skills necessary to tackle more complex and higher-order kinds of instructional tasks that require problem-solving and critical thinking. Teachers' reported that students need to develop more sophisticated study skills than simply memorizing information, yet the largest percentage of items that appeared on their tests and other supporting documents were categorized as encoding, calling for basic knowledge of content. Students may not engage in more advanced kinds of study skills because the course exams and other assignments simply do not demand it. Other research has shown that studying and achievement is influenced by student expectations about the challenge or demand associated with criterion performance (Hamaker, 1986; Martin & Saljo, 1976; Rothkopf & Billington, 1975;

Thomas, et al., in press). Teachers may verbalize the need for students to develop more sophisticated study strategies but do not provide the demands and practice that would promote this development.

The findings not only indicate that assessment items are not very challenging in terms of cognitive demand, but that teachers frequently include items identical to those that have appeared on other course documents, and that they are not aware they are engaging in this type of compensatory practice. Semb and Spencer (1976) also found that instructors did not accurately judge the complexity of test items and over-estimated the percentage of problem-solving items that appeared on their exams. Because compensations reduce the complexity and demand of test items and are associated with less productive types of study strategies (Thomas et al., in press), an awareness that these types of items are included on exams would help teachers increase the challenge of their course exams.

To further illustrate the relationships between achievement goals, the types of study skills students need to improve and assessment practices, the pattern of findings across variables for one teacher will be described. Teacher 5 said she wanted students in her course to learn about themselves, the real-world issues that affected them, and to develop an enthusiasm for biology. She hoped students would understand "environmental issues, genetics, ethics and other societal concerns" informed by biology and science. In terms of the study skills she wanted her students to develop, she identified "problem-solving, perceiving inter-relationships, and moving from facts and recall to interpretation". Clearly these are admirable goals and skills but an examination of her test items, support items, and compensatory practices paint a very different picture. Even though she isolated problem-solving as a study skill that students needed to develop, there were no extension items on her exam and only 3 percent of the items on all other course documents required application. Fully 53 percent of her test items and 69 percent of her practice items were categorized as encoding and do not seem to reinforce her hope that students would move from recalling facts to interpretation. Moreover, she estimated that only two percent of her test items were ones that students had seen previously, when, in fact, 64 percent of her test items were items that previously

appeared on other course materials. Undeniably her intentions for student achievement and study skill development were exemplary, and in all likelihood, she was simply unaware of the misalignment between her instructional goals and her assessment practices.

One question that remains to be addressed is how to promote greater congruence between instructional goals and assessment practices. How are the good intentions of teachers, like those of Teacher 5, translated into good assessment practices that support their achievement and study skill development goals? The results obtained in this study suggest that teachers' may be better able to meet their professed learning objectives if they utilize some type of blueprint for classifying individual items. Other researchers have also recommended that teachers adopt some type of classification scheme for item generation and have published their classification schemes (e.g., Biggs & Collis, 1982; Bloom, et al., 1956). A very simple blueprint based on the scheme used to categorize items in this study that might be employed by teachers is provided in Figure 1.

Figure 1. Blueprint for Item Development.

Production	Level of Processing			Compensation
	Encoding	Integration	Extension	(Identical item?)
True/ false				
Multiple choice or matching				
Fill in the blanks				
Short answer				
Essay/ complex production				

Teachers would be able to select or develop items categorized by production type, level of processing and also note whether the students have seen an identical item previously (a compensation). This type of blueprint might encourage teachers to evaluate the kinds of cognitive processes demanded on their tests and other course assignments, and help teachers build assessment instruments that meet their stated goals and objectives.

One issue that remains unresolved is whether teachers emphasize low-level items on their tests because they implicitly assume they are risking student failure if their tests require critical thinking skills like application and problem-solving. In other words, the teachers may have identified the kinds of study skills they think students need to develop but perceive that their students will be unsuccessful on items calling for higher level types of strategies and skills. This may be one reason why the compensatory practice of including identity items on the test was so prevalent. One practical suggestion to help solve this dilemma is to support engagement in more advanced kinds of cognitive processes and study strategies by providing practice items that establish expectations and encourage student engagement in higher order thinking skills. Compensations may be minimized by without risking widespread student failure by changing the nature of compensations so that they act as supports. For instance, instead of providing students with exact replicas of items that appear on the test, sample items and review questions that require high levels of processing and production could be provided. Supportive practices may prove to be a useful strategy for promoting critical thinking and realizing course goals.

REFERENCES

Bereiter, C., & Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher-order thinking skills in reading and writing. *Curriculum Inquiry, 17*, 9-30.

Berliner, D. C. (1992). Redesigning classroom activities for the future. *Educational Technology, 32*, 7-13.

Biggs, J. B, & Collis, K .F. (1982). *Evaluating the quality of learning: The SOLO Taxonomy.* New York: Academic Press.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The cognitive domain.* New York: McKay

Bol, L., & Thomas, J. W. (1991). *The relationship between teachers' classroom practices and students' study activities in high-school Biology courses.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438-481.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*, 159-199.

Fleming, M., & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W. E. Hathaway (Ed.), *New directions for testing and measurement: Vol. 19. Testing in the schools* (pp. 29-38). San Francisco: Josey Bass.

Haertel, E. (1986). *Choosing and using classroom tests: Teachers perspectives on assessment.* Paper presented at the annual Meeting of the American Educational Research Association, San Francisco.

Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research, 56*, 212-242.

Linn, R. L. (1990). Essentials of student assessment: From accountability to instructional aid. In S. Tozer, T.H. Anderson, & B.B. Armbruster (Eds.), *Foundational Studies in Teacher Education: A Reexamination* (pp. 124-138). New York: Teachers College Press.

Madaus, G. F., Maxwell West, M., Harmon, M. C., Lomax, R. G., & Viator, K. A. (1992). *The influence of testing on teaching math and science in grades 4-12: Executive summary.* Center for the Study of Testing, Evaluation, and Educational Policy, Boston College: National Science Foundation.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning: II. Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology, 46*, 115-127.

Mayer, R. E. (1979). Can advance organizers influence meaningful learning? *Review of Educational Research, 49*, 371-383.

McCaslin, M., & Good, T. L. (1992). Compliant cognition: The misalliance of management and instructional goals in current school reform. *Educational Researcher, 21*, 4 - 17.

Meyer, G. (1936). The effects on recall and recognition of the examination set in classroom situations. *Journal of Educational Psychology, 27,* 81-99.

Newmann, F.M. (1990). Higher-order thinking in teaching social studies: A rationale for the assessment of classroom thoughtfulness. *Journal of Curriculum Studies, 22*(1), 41-56.

Peterson, P. L. (1988). Teaching for higher-order thinking in mathematics: The challenge for the next decade. In D.A. Grouws, T.J. Cooney, & D.Jones (Eds.), *Effective Mathematics Teaching* (pp. 2-26). National Council of Teachers of Mathematics: Erlbuam.

Prawat, R. S. (1993). The value of ideas: Problems versus possibilities in learning. *Educational Researcher, 22*(6), 5-16.

Rothkopf, E. Z., & Billington, M. J. (1975). A two-factor model of the effect of goal-descriptive directions on learning from text. *Journal of Educational Psychology, 67,* 692-704.

Semb, G., & Spencer, R. (1976). Beyond the level of recall: An analysis of complex educational tasks in college and university instruction. In L.E. Farley and E.A. Vargus (Eds.), *Behavioral Research and Technology* (pp. 115-125). Procedings of 3rd National Conference.

Strage, A., Tyler, A. B., Thomas, J. W., & Rohwer, W. D., Jr. (1987). An analytic framework for assessing distinctive course features within and across grade levels. *Contemporary Educational Psychology, 12,* 280-302.

Thomas, J.W., Bol, L., & Warkentin, R.W. (1991). Antecedents of college students' study deficiencies: The relationship between course features and students' study activities. *Higher Education, 22,* 275-296.

Thomas, J.W., Bol, L., Warkentin, R.W., Wilson, M., Strage, A., & Rohwer, W.D., (in press). Interrelationships among, students' study activities, self-concept of academic ability, and achievement in high school biology courses. *Applied Cognitive Psychology.*